

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/653,681B  
Source: TFW16  
Date Processed by STIC: 7/14/05

***ENTERED***

~~CRF ERRORS EDITED BY THE STIC SYSTEMS  
Branch~~

Serial Number: 10/653,681B

CRF Edit Date:

Edited by:

7/15/05

Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

Corrected the SEQ ID NO. Sequence numbers edited were:

✓ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

5

Deleted: \_\_\_ invalid beginning/end-of-file text ; \_\_\_ page numbers

Inserted mandatory headings/numeric identifiers, specifically:

Moved responses to same line as heading/numeric identifier, specifically:

Other:

Sequence 5 - reduced blank spaces between groups  
of 10 bases



IFW16

RAW SEQUENCE LISTING DATE: 07/15/2005  
 PATENT APPLICATION: US/10/653,681B TIME: 12:24:17

Input Set : A:\PTO.AMC.txt  
 Output Set: N:\CRF4\07152005\J653681B.raw

```

3 <110> APPLICANT: DAI, KEN-SHWO
5 <120> TITLE OF INVENTION: HUMAN ARL-RELATED GENE VARIANTS ASSOCIATED WITH CANCER
7 <130> FILE REFERENCE: U014798-3
9 <140> CURRENT APPLICATION NUMBER: 10/653,681B
10 <141> CURRENT FILING DATE: 2003-09-02
12 <160> NUMBER OF SEQ ID NOS: 6
14 <170> SOFTWARE: PatentIn version 3.2
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 1090
18 <212> TYPE: DNA
19 <213> ORGANISM: ARTIFICIAL SEQUENCE
21 <220> FEATURE:
22 <223> OTHER INFORMATION: VARIANT OF HUMAN ALDOSE REDUCTASE-LIKE GENE
25 <220> FEATURE:
26 <221> NAME/KEY: CDS
27 <222> LOCATION: (70)..(333)
29 <400> SEQUENCE: 1
30 caaaaacagc aacagaaaagc aggacgtgag acttctacct gctcaactcg aatcatttct 60
32 gcaccaacc atg gcc acg ttt gtg gag ctc agt acc aaa gcc aag atg ccc 111
33 Met Ala Thr Phe Val Leu Ser Thr Lys Ala Lys Met Pro
34 1 5 10
36 att gtg ggc ctg ggc act tgg aag tct cct ctc ggc aaa gtg aaa gaa 159
37 Ile Val Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu
38 15 20 25 30
40 gca gtg aag gtg gcc att gat gca gga tat cgg cac att gac tgt gcc 207
41 Ala Val Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala
42 35 40 45
44 tat gtc tat cag aat gaa cat gaa gtg ggg gaa gcc atc caa gag aag 255
45 Tyr Val Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys
46 50 55 60
48 atc caa gag aag gct gtg aag cgg gag gac ctg ttc atc gtc agc aag 303
49 Ile Gln Glu Lys Ala Val Lys Arg Glu Asp Leu Phe Ile Val Ser Lys
50 65 70 75
52 ttg tgg ccc act tcc aga tcg aga agc tct tgaacaaacc tggactgaaa 353
53 Leu Trp Pro Thr Ser Arg Ser Arg Ser Ser
54 80 85
56 tataaaccag tgactaacca ggttgagtgt cacccataacc tcacgcaggaa gaaactgatc 413
58 cagtaactgcc actccaagggtt catcacccgtt acggcctaca gccccctggg ctctccggat 473
60 agaccttggg ccaaggccaga agacccttcc ctgctggagg atcccaagat taaggagatt 533
62 gctgcaaagc aaaaaaaaaac cgccggccag gttctgatcc gtttccatat ccagaggaat 593
64 gtgattgtca tccccaaagtc tgtgacacca gcacgcattt ttgagaacat tcaggtcttt 653
66 gactttaaat tgagtgtatga ggagatggca accataactca gcttcaacag aaactggagg 713
68 gcctgtaaacg tggatgtatca ctctcatttgc gaagactatc ctttcgtatgc agaatattga 773

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/653,681B

DATE: 07/15/2005

TIME: 12:24:17

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\07152005\J653681B.raw

70 ggttgaatct cctggtgaga ttatacagga gattctcttt ctgcgtgaa gtgtgactac 833  
 72 ctccactcat gtccccattt agccaagctt atttaagatc acagtgaact tagtcctgtt 893  
 74 atagacgaga atcgagggtgc tggttagac atttatttct gtatgttcaa ctaggatcag 953  
 76 aatatcacag aaaagcatgg ctgaaataag gaaatgacaa tttttccac ttatctgatc 1013  
 78 agaacaatg tttatataagc atcagaaaact ctgccaacac tgaggatgt aagatcaata 1073  
 80 aaaaaataa taatcat 1090  
 83 <210> SEQ ID NO: 2  
 84 <211> LENGTH: 88  
 85 <212> TYPE: PRT  
 86 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 88 <220> FEATURE:  
 89 <223> OTHER INFORMATION: PEPTIDE ENCODED BY VARIANT OF HUMAN ALDOSE REDUCTASE-LIKE

GENE

91 <400> SEQUENCE: 2  
 93 Met Ala Thr Phe Val Glu Leu Ser Thr Lys Ala Lys Met Pro Ile Val  
 94 1 5 10 15  
 97 Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu Ala Val  
 98 20 25 30  
 101 Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala Tyr Val  
 102 35 40 45  
 105 Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys Ile Gln  
 106 50 55 60  
 109 Glu Lys Ala Val Lys Arg Glu Asp Leu Phe Ile Val Ser Lys Leu Trp  
 110 65 70 75 80  
 113 Pro Thr Ser Arg Ser Arg Ser Ser  
 114 85

117 <210> SEQ ID NO: 3  
 118 <211> LENGTH: 1279  
 119 <212> TYPE: DNA  
 120 <213> ORGANISM: ARTIFICIAL SEQUENCE  
 122 <220> FEATURE:  
 123 <223> OTHER INFORMATION: VARIANT OF HUMAN ALDOSE REDUCTASE-LIKE GENE  
 126 <220> FEATURE:  
 127 <221> NAME/KEY: CDS  
 128 <222> LOCATION: (70)..(804)  
 130 <400> SEQUENCE: 3

131 caaaaacagc aacagaaaagc aggacgtgag acttctacct gctcaactcag aatcatttct 60  
 133 gcaccaacc atg gcc acg ttt gtg gag ctc agt acc aaa gcc aag atg ccc 111  
 134 Met Ala Thr Phe Val Glu Leu Ser Thr Lys Ala Lys Met Pro  
 135 1 5 10  
 137 att gtg ggc ctg ggc act tgg aag tct ctc ggc aaa gtg aaa gaa 159  
 138 Ile Val Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu  
 139 15 20 25 30  
 141 gca gtg aag gtg gcc att gat gca gga tat cgg cac att gac tgt gcc 207  
 142 Ala Val Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala  
 143 35 40 45  
 145 tat gtc tat cag aat gaa cat gaa gtg ggg gaa gcc atc caa gag aag 255  
 146 Tyr Val Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys  
 147 50 55 60  
 149 atc caa gag aag gct gtg aag cgg gag gac ctg ttc atc gtc agc aag 303

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/653,681B

DATE: 07/15/2005

TIME: 12:24:17

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\07152005\J653681B.raw

150	Ile	Gln	Glu	Lys	Ala	Val	Lys	Arg	Glu	Asp	Leu	Phe	Ile	Val	Ser	Lys
151	65						70						75			
153	tgg	tgg	ccc	act	ttc	ttt	gag	aga	ccc	ctt	gtg	agg	aaa	gcc	ttt	gag
154	Leu	Trp	Pro	Thr	Phe	Phe	Glu	Arg	Pro	Leu	Val	Arg	Lys	Ala	Phe	Glu
155	80						85					90				
157	aag	acc	ctc	aag	gac	ctg	aag	ctg	agc	tat	ctg	gac	gtc	tat	ctt	att
158	Lys	Thr	Leu	Lys	Asp	Leu	Lys	Leu	Ser	Tyr	Leu	Asp	Val	Tyr	Leu	Ile
159	95						100				105			110		
161	cac	tgg	cca	cag	gga	ttc	aag	tct	ggg	gat	gac	ctt	ttc	ccc	aaa	gat
162	His	Trp	Pro	Gln	Gly	Phe	Lys	Ser	Gly	Asp	Asp	Leu	Phe	Pro	Lys	Asp
163	115							120				125				
165	gat	aaa	ggt	aat	gcc	atc	ggt	gga	aaa	gca	acg	ttc	ttg	gat	gcc	tgg
166	Asp	Lys	Gly	Asn	Ala	Ile	Gly	Gly	Lys	Ala	Thr	Phe	Leu	Asp	Ala	Trp
167	130							135				140				
169	gag	gcc	atg	gag	gag	ctg	gtg	gat	gag	ggg	ctg	gtg	aaa	gcc	ctt	ggg
170	Glu	Ala	Met	Glu	Glu	Leu	Val	Asp	Glu	Gly	Leu	Val	Lys	Ala	Leu	Gly
171	145						150				155					
173	gtc	tcc	aat	ttc	agc	cac	ttc	cag	atc	gag	aag	ctc	ttg	aac	aaa	cct
174	Val	Ser	Asn	Phe	Ser	His	Phe	Gln	Ile	Glu	Lys	Leu	Leu	Asn	Lys	Pro
175	160						165				170					
177	gga	ctg	aaa	tat	aaa	cca	gtg	act	aac	cag	gtt	gag	tgt	cac	cca	tac
178	Gly	Leu	Lys	Tyr	Lys	Pro	Val	Thr	Asn	Gln	Val	Glu	Cys	His	Pro	Tyr
179	175						180				185			190		
181	ctc	acg	cag	gag	aaa	ctg	atc	cag	tac	tgc	cac	tcc	aag	ggc	atc	acc
182	Leu	Thr	Gln	Glu	Lys	Leu	Ile	Gln	Tyr	Cys	His	Ser	Lys	Gly	Ile	Thr
183	195							200				205				
185	gtt	acg	gcc	tac	agc	ccc	ctg	ggc	tct	ccg	gat	aga	cct	tgg	gcc	aag
186	Val	Thr	Ala	Tyr	Ser	Pro	Leu	Gly	Ser	Pro	Asp	Arg	Pro	Trp	Ala	Lys
187	210							215				220				
189	cca	gaa	gac	cct	tcc	ctg	ctg	gag	gat	ccc	aag	att	aag	gag	att	gct
190	Pro	Glu	Asp	Pro	Ser	Leu	Leu	Glu	Asp	Pro	Lys	Ile	Lys	Glu	Ile	Ala
191	225						230				235					
193	gca	aag	cac	tcc	cca	agt	ctg	tgacaccaggc	acgcattgtt	gagaacattc					834	
194	Ala	Lys	His	Ser	Pro	Ser	Leu									
195	240						245									
197	agg	tct	ttt	ga	ctt	taa	att	tg	at	gg	ca	ac	cata	tc	ag	aa
199	act	gg	gg	gc	ct	gt	taa	c	ac	gt	tt	cc	tat	tc	ca	ga
201	aat	ttt	gagg	tt	gaat	tc	cc	t	tt	gt	gg	at	ca	gg	at	tt
203	gt	act	ac	ct	cc	act	cat	gt	cc	at	tt	tt	tt	tt	tc	tt
205	gt	cct	gt	tt	at	aa	cg	ag	gt	gt	tt	tt	tt	tt	tt	tt
207	agg	atc	agg	aa	tat	cac	agg	gg	t	gaata	agg	aa	tat	gac	aa	t
209	at	ct	gt	ca	a	aca	at	gt	tat	taa	agg	cat	ca	gaa	act	tg
211	gt	ca	at	aa	aaa	aaa	ata	ta	at	at	ca	ttt	cc	act	tg	aaa
214	<210>	SEQ	ID	NO:	4											
215	<211>	LENGTH:	245													
216	<212>	TYPE:	PRT													
217	<213>	ORGANISM:	ARTIFICIAL	SEQUENCE												
219	<220>	FEATURE:														
220	<223>	OTHER INFORMATION:	PEPTIDE	ENCODED	BY	VARIANT	OF	HUMAN	ALDOSE	REDUCTASE-LIKE						
GENE																

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/653,681B

DATE: 07/15/2005

TIME: 12:24:17

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\07152005\J653681B.raw

222 <400> SEQUENCE: 4  
 224 Met Ala Thr Phe Val Glu Leu Ser Thr Lys Ala Lys Met Pro Ile Val  
 225 1 5 10 15  
 228 Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu Ala Val  
 229 20 25 30  
 232 Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala Tyr Val  
 233 35 40 45  
 236 Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys Ile Gln  
 237 50 55 60  
 240 Glu Lys Ala Val Lys Arg Glu Asp Leu Phe Ile Val Ser Lys Leu Trp  
 241 65 70 75 80  
 244 Pro Thr Phe Phe Glu Arg Pro Leu Val Arg Lys Ala Phe Glu Lys Thr  
 245 85 90 95  
 248 Leu Lys Asp Leu Lys Leu Ser Tyr Leu Asp Val Tyr Leu Ile His Trp  
 249 100 105 110  
 252 Pro Gln Gly Phe Lys Ser Gly Asp Asp Leu Phe Pro Lys Asp Asp Lys  
 253 115 120 125  
 256 Gly Asn Ala Ile Gly Gly Lys Ala Thr Phe Leu Asp Ala Trp Glu Ala  
 257 130 135 140  
 260 Met Glu Glu Leu Val Asp Glu Gly Leu Val Lys Ala Leu Gly Val Ser  
 261 145 150 155 160  
 264 Asn Phe Ser His Phe Gln Ile Glu Lys Leu Leu Asn Lys Pro Gly Leu  
 265 165 170 175  
 268 Lys Tyr Lys Pro Val Thr Asn Gln Val Glu Cys His Pro Tyr Leu Thr  
 269 180 185 190  
 272 Gln Glu Lys Leu Ile Gln Tyr Cys His Ser Lys Gly Ile Thr Val Thr  
 273 195 200 205  
 276 Ala Tyr Ser Pro Leu Gly Ser Pro Asp Arg Pro Trp Ala Lys Pro Glu  
 277 210 215 220  
 280 Asp Pro Ser Leu Leu Glu Asp Pro Lys Ile Lys Glu Ile Ala Ala Lys  
 281 225 230 235 240  
 284 His Ser Pro Ser Leu  
 285 245  
 287 <210> SEQ ID NO: 5  
 289 <211> LENGTH: 1337  
 291 <212> TYPE: DNA  
 293 <213> ORGANISM: HOMO SAPIEN  
 295 <220> FEATURE:  
 297 <223> OTHER INFORMATION: HUMAN ARL GENE  
 299 <400> SEQUENCE: 5  
 301 caaaaacagc aacagaaaagc aggacgtgag acttctaccgt gctcaactcg aatcattttct 60  
 302 gcaccaaccca tggccacgtt tggggagctc agtaccaaag ccaagatgcc catttgtggc 120  
 303 ctgggcactt ggaagtctcc tctcgccaaa gtgaaaagaag cagtgaaggt ggccattgtat 180  
 304 gcaggatatac ggcacattga ctgtgcctat gtctatcaga atgaacatga agtggggaa 240  
 305 gccatccaag agaagatcca agagaaggct gtgaagcggg aggacctgtt catcgtcagc 300  
 306 aagttgtggc ccactttctt tgagagaccc cttgtgagga aagccttga gaagaccctc 360  
 307 aaggacctga agctgagcta tctggacgtc tatcttattc actggccaca gggattcaag 420  
 308 tctggggatg acctttccc caaagatgat aaaggtaatg ccatcggtgg aaaagcaacg 480  
 309 ttcttgatg cctggagggc catggaggag ctgggtggatg aggggctggt gaaagccctt 540

## **RAW SEQUENCE LISTING**

PATENT APPLICATION: US/10/653,681B

DATE: 07/15/2005

TIME: 12:24:17

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\07152005\J653681B.raw

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/653,681B

DATE: 07/15/2005

TIME: 12:24:19

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\07152005\J653681B.raw

**Raw Sequence Listing before editing,  
for reference only**



IFW16

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/653,681B

DATE: 07/14/2005

TIME: 12:29:14

Input Set : A:\U0147983 SEQUENCE LISTING.txt

Output Set: N:\CRF4\07142005\J653681B.raw

3 <110> APPLICANT: DAI, KEN-SHWO  
 5 <120> TITLE OF INVENTION: HUMAN ARL-RELATED GENE VARIANTS ASSOCIATED WITH CANCER  
 7 <130> FILE REFERENCE: U014798-3  
 9 <140> CURRENT APPLICATION NUMBER: 10/653,681B  
 10 <141> CURRENT FILING DATE: 2003-09-02  
 12 <160> NUMBER OF SEQ ID NOS: 6  
 14 <170> SOFTWARE: PatentIn version 3.2

## ERRORED SEQUENCES

pp 13

Does Not Comply  
with  
recommended Diskette Needle

287 <210> SEQ ID NO: 5						
289 <211> LENGTH: 1337						
291 <212> TYPE: DNA						
293 <213> ORGANISM: HOMO SAPIEN						
295 <220> FEATURE:						
297 <223> OTHER INFORMATION: HUMAN ARL GENE						
299 <400> SEQUENCE: 5						
301 caaaaacagc	← delete extra spaces - per sequence Rule, only one space between groups of 10	aacagaaaagc	aggacgtgag	acttctacct	gctcactcag	aatcatttct
302 gcaccaacca	tggccacgtt	tgtggagctc	agtaccaaag	ccaagatgcc	cattgtggc	120
303 ctggcactt	ggaagtctcc	tctcgccaaa	gtgaaagaag	cagtgaaggt	ggccatttgat	180
304 gcaggatatc	ggcacattga	ctgtgcctat	gtctatcaga	atgaacatga	agtggggaa	240
305 gccatccaag	agaagatcca	agagaaggct	gtgaagcggg	aggacctgtt	catcgtcagc	300
306 aagttgtggc	ccactttctt	tgagagaccc	cttgtgagga	aagccttga	gaagaccctc	360
307 aaggaccta	agctgagcta	tctggacgtc	tatcttattc	actggccaca	gggattcaag	420
308 tctggggatg	accttttccc	caaagatgat	aaaggtaatg	ccatcggtgg	aaaagcaacg	480
309 ttcttggatg	cctggggaggc	catggaggag	ctgggtggatg	aggggctgg	gaaagccctt	540
310 ggggtctcca	atttcagcca	cttccagatc	gagaagctct	tgaacaaacc	tggactgaaa	600
311 tataaaccag	tgactaacca	ggttgagtgt	cacccataacc	tcacgcagga	gaaactgatc	660
312 cagtactgcc	actccaaggg	catcaccgtt	acggcctaca	gccccctggg	ctctccggat	720
313 agaccttggg	ccaagccaga	agacccttcc	ctgctggagg	atcccaagat	taaggagatt	780

314						
gctgcaaagc	acaaaaaaaaac	cgcagcccaag	gttctgatcc	gtttccatat	ccagaggaat	840
315						
tgattgtca	tccccaaagtc	tgtgacaccca	gcacgcattg	ttgagaacat	tcaggtcttt	900
316						
gactttaaat	tgagtgtatga	ggagatggca	accatactca	gcttcaacag	aaactggagg	960
317						
gcctgtaacg	tgttgcaatc	ctctcatttg	gaagactatc	ccttcgatgc	agaatattga	1020
318						
ggttgaatct	cctgggtgaga	ttatacagga	gattctcttt	cttcgctgaa	gtgtgactac	1080
319						
ctccactcat	gtcccatttt	agccaaagctt	attnaagatc	acagtgaact	tagtcctgtt	1140
320						
atagacgaga	atcgagggtgc	tgttttagac	atttattttct	gtatgttcaa	ctaggatcag	1200
321						
aatatcacag	aaaagcatgg	cttgaataag	gaaatgacaa	ttttttccac	ttatctgatc	1260
322						
agaacaaatg	tttattaagc	atcagaaaact	ctgccaacac	tgaggatgta	aagatcaata	1320
E--> 323						

aaaaaaataa taatcat

*sane error*

*sel P.3*

10/653, 681B 3

aagatcaata 1320

taatcat

1380

1337

aaaaaaaataa

RAW SEQUENCE LISTING ERROR SUMMARY                    DATE: 07/14/2005  
PATENT APPLICATION: US/10/653,681B                    TIME: 12:29:15

Input Set : A:\U0147983 SEQUENCE LISTING.txt  
Output Set: N:\CRF4\07142005\J653681B.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:5; Line(s) 301,302,303,304,305,306,307,308,309,310,311,312,313,314,315  
Seq#:5; Line(s) 316,317,318,319,320,321,322,323

**VERIFICATION SUMMARY** DATE: 07/14/2005  
PATENT APPLICATION: US/10/653,681B TIME: 12:29:15

Input Set : A:\U0147983 SEQUENCE LISTING.txt  
Output Set: N:\CRF4\07142005\J653681B.raw

L:323 M:254 E: No. of Bases conflict, LENGTH:Input:1380 Counted:1337 SEQ:5